**Graphical Representation of Data**

### **Data Patterns in Statistics**

In general, we can define the patterns in data as follows:

center, spread, shape, and unusual features.

Some known distributions have some special descriptive labels such as symmetric, bell-shaped, skewed, etc.

#### **1-Center**

The center of a distribution, graphically, is located at the median of the distribution.

Chart, histogram

Description automatically generated

#### **2-Spread**

If the set of observations spreads over a wide range, the spread is greater. If the observations are centered around a single value in a narrower range, then the spread is smaller.

Chart, histogram

Description automatically generated

#### **3-Shape**

**Symmetry**

In the Normal Distribution, the two sides of the distribution are equal and symmetrical.

Chart, histogram

Description automatically generated

***Number of peaks***

Distributions with one or multiple peaks.

Distribution with one clear peak is known as unimodal, and distribution with two clear peaks is called bimodal.

A single peak symmetric distribution at the center, is referred to as bell-shaped.

Chart, histogram

Description automatically generated

***Skewness***

Some distributions may have multiple observations on one side of the graph than the other side.

Distributions having fewer observations towards higher values are said to be skewed right

Distributions with fewer observations towards lower values are said to be skewed left.

Chart, histogram

Description automatically generated

***Uniform***

When the set of observations has no peak and have data equally spread across the range of the distribution

Chart, histogram

Description automatically generated

#### **4-Unusual Features**

***Gaps***

Gaps points to areas of a distribution having no observations.

Following figure has a gap as there are no observations in the middle of the distribution.

Chart, bar chart, histogram

Description automatically generated

***Outliers***

Distributions may be characterized by extreme values that differ greatly from the other set of observation data

Chart, histogram

Description automatically generated

### **Frequency Table**

A **frequency table** is a table that lists items and shows the number of times the items occur. The relative frequency of a category or a numerical value is the associated frequency divided by the total number of data.

Table

Description automatically generated

### **Pie Chart**

A pie chart is a circle having a “slice of the pie” for each category.

Chart, pie chart

Description automatically generated

Pie charts are effective for displaying the relative frequencies of a small number of categories

They are not recommended, however, when you have a large number of categories.

Pie charts can also be confusing when they are used to compare the outcomes of two different surveys or experiments.

When slices become too small, pie charts have to rely on colors, textures or arrows so the reader can understand them.

**Bar Chart**

Bar charts can be used to represent frequencies of different categories.

Typically, the Y-axis shows the number of observations in each category rather than the percentage of observations as is typical in pie charts.

A vertical bar chart is sometimes called a column chart.

Chart, bar chart

Description automatically generated

Histogram

A histogram is a graphical method for displaying the shape of a distribution. It is particularly useful when there are a large number of observations.

To create the table, the range of scores can broken into intervals, called class intervals.

The widths of the class intervals, sometimes called bin widths. Your choice of bin width determines the number of class intervals.

**Example:**

250 students on a statistics test.

The students' scores ranged from 32 to 96.

A simple frequency table would be too big, containing over 100 rows.

To simplify the table, we group scores together as shown in table below.

Table

Description automatically generatedChart, histogram

Description automatically generated

With bar charts, the labels on the X axis are categorical; with histograms, the labels are quantitative.

**Difference Between Histogram and Bar Graph**

That there are gaps between bars in a bar graph(Bar Chart) but in the histogram, the bars are adjacent to each other.

Chart, histogram

Description automatically generated

|  |  |
| --- | --- |
| compare different categories of data(Spanish, Chinese…) | One category(GDP)  frequency of numerical data. |
| Comparison of discrete variables | Distribution of non-discrete variables |
| Categorical data | Quantitative data |
| there are spaces between bars. | there are no spaces between bars |
| You can rearrange the blocks, from highest to lowest | They are shown in the sequences of classes |
| Width of bars are same | Need not to be same |
| The width of the bars in a  bar graph is always same. | The width of rectangular blocks in a histogram may or may not be same |